CENTER FOR CURRICULUM REDESIGN

Mathematics for the 21st Century

Paper #2

WHAT should students learn?

Methods and tools

By Sanjoy Mahajan, PhD

Massachusetts Institute of Technology, Olin College of Engineering *(present)*

University of Cambridge, Caltech, University of Oxford *(previous)*

July 23, 2014

With sincere thanks, and appreciation for their generous support, to the:

BILL & MELINDA GATES foundation

Fondation Henri Moser

www.curriculumredesign.org

Summary

Understanding mathematics is essential for full participation in society. Yet mathematics is learned mostly by rote, with little understanding or possibility for transfer. Thus, PISA can provoke a revolution in mathematics education worldwide---by focusing the mathematics assessment on distinguishing rote learning from true understanding. Assessment drives teaching and learning: What PISA measures, the mathematics curriculum worldwide will deliver.

Introduction

Almost 30 years ago, one of the authors (SM) last helped a fellow student cheat. The class was civics. Our teacher, former classmate of Apollo-11 pilot Buzz Aldrin, wondering whether we could find Europe on a map, sprang upon us with a geography test. Its hardest question asked how far several countries were from the borders of civilization. My neighbor in the first row of desks whispered: "How far is Cuba from the United States? A thousand miles?" I whispered back, "No, ninety!" The next day, Buzz Aldrin's schoolmate returned our tests. We knew, he said, more geography than he had expected. However, one student placed Cuba 90,000 miles from the United States!

A comparison with other earthly distances---for example, the 3000 miles across the entire continental United States---would have shown my class neighbor that, when forced to choose between 90 and 90,000 miles, the correct distance had to be 90 miles. However, this powerful reasoning tool of comparison had not been part of his mathematical education. Handling the numbers without understanding or meaning, he picked a distance almost halfway to the moon.

He lacked a powerful ability: using numbers to understand and make sense of the world. If in your mind Cuba could as well be halfway to the moon, numbers are not helping you understand the physical layout of the world. Columbus, despite all his geography mistakes, had a more accurate picture of the world 500 years ago.

Mathematical understanding is crucial for high performance in our personal, public, and work lives. At home, we may want to understand the results of a medical test, or rekindle our child's love of math. At work, we may need to estimate the money, time, and employees for a large project. As citizens, we may want to judge the rise in carbon-dioxide levels in the air, or the proportion of tax dollars that should go to health, education, or war. Finally, mathematics underlies our science, technology, and engineering.

A society where a small elite has the legitimacy to govern has little need for mathematics as a social language for resolving questions. The members of the elite share a worldview and trust. In democratic societies, the elite is diverse, the shared trust is absent, and it is replaced by numbers---hence the prevalence of cost-benefit analyses.[1] Indeed, geometry and deductive argument arose in ancient Athens because of the need for them in the democratic Athenian courts and juries.[2] Mathematics became the only universal and impartial language, and understanding mathematics becomes essential for "full participation in society".

PISA's goal is widely applicable real learning

1 Theodore Porter (1995), *Trust in Numbers*, Princeton University Press.

2 G.E.R. Lloyd (1990), *Demystifying Mentalities*, Cambridge University Press; G.E.R. Lloyd (1992), "Democracy, philosophy, and science in ancient Greece," in John Dunn (ed.), *Democracy: The Unfinished Journey, 508 BC to 1993*, Oxford University Press, pp. 41-56.

The breadth of mathematical application is reflected in the four contexts assessed by PISA:[3]

1. personal: self, family and peer groups
2. societal: one's community
3. occupational: the world of work
4. scientific: application to science and related issues and topics

These contexts are outstanding choices. Furthermore, by weighting them equally, PISA ameliorates the misconception that mathematics is useful only in the scientific context.

Furthermore, PISA aims to assess whether students can “extrapolate from what they know,” not simply reproduce knowledge.[4] Without making meaning and sense of mathematics, students cannot extrapolate or transfer their knowledge to new domains. Rote knowledge---learning without understanding or meaning---is inert. It is the fundamental obstacle to PISA's goal, that mathematics help prepare students for full participation in society.

Origin of rote learning

It is understandable that the traditional mathematics curriculum would produce rote learning, because our mathematics curriculum grew up in an era with different needs. At first, arithmetic was necessary for monetary calculations needed by the rising merchant class and Fibonacci's *Liber Abaci* (1202) was the manual of instruction.[5] Even through much of the twentieth century, if an exact calculation was needed, a human had to do it. The Enigma codebreaking effort and the Manhattan project even relied on human computers. Pressed by this need, we developed precise algorithms for long multiplication and division, even for extracting square roots, and in mathematics students learned these procedures.

However, with the advent of electronic calculators, and computational technology in general, the human and the computer could specialize. In the 21st century, calculators and computers can make the tedious calculations. Humans will keep the rest: *making understanding and meaning*. Mathematics teaching can and should change accordingly.

The contrast between human and computer thinking is illustrated in an old experiment by the Gestalt psychologist Max Wertheimer, who describes observing students learning the fundamental arithmetic operations.[6] The students were solving the problems set by the teacher, seemingly understanding what they were doing. Curious about the depth of understanding, Wertheimer gave the class the following question:

$$\frac{243 + 243 + 243 + 243}{4} = ?$$

There are three kinds of response: laughter, despair ("That's too hard."), and computation. The best answer is laughter; these students understood the connection between addition and multiplication. The worst answer is to do the calculation. This answer reveals that the students, even if they could

---

3 OECD (2013), PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy, OECD Publishing, p. 37 (p. 38 in the PDF file).

4 Andreas Schleicher (2012), "Use data to build better schools", TEDGlobal 2012.

5 Keith Devlin (2011), The Man of Numbers: Fibonacci's Arithmetic Revolution, Walker.

6 Max Wertheimer (1945/1959), Productive Thinking (Michael Wertheimer, ed.), Harper.

do arithmetic manipulations with great skill, did not have any clue what the arithmetic operations meant.

Imagine trying Wertheimer's experiment with computers as the students. The computers cannot laugh. Nor can the computers throw up their hands in despair. Rather, the computers would simply compute correctly. The natural computer method is the worst human method. The best human answer, laughter, is beyond the comprehension of any computer. Computer and human reasoning are opposites. The kind of learning that humans and computers need---real and rote learning, respectively---are opposites.

Rote learning is the fundamental obstacle

In all four PISA contexts, the fundamental barrier to high performance is rote or parrot learning. Davydov distinguishes rote versus real learning as empirical versus theoretical thinking, with empirical thinking (rote learning) focusing on the surface features of a problem.[7] As a result, rote learning is far less efficient than real learning.

The classic comparison is Galperin and Pantina's study of teaching children to write Russian characters.[8] In comparison to children who simply practiced from model letters, children taught a theoretical understanding---the underlying structure of the letter strokes and how to describe them---learned with far fewer repetitions and greater accuracy and transfer to new domains (such as copying blueprints). The lesson for mathematics education is that whenever students need a great deal of drill, it is likely that they are learning by rote.

The canonical example of rote learning comes from the third National Assessment of Education Progress (NAEP) in the United States, students performed exact long multiplication with high success rates (for 13- and 17-year olds, they were 57- and 72 percent, respectively).[9] However, when students were asked to "Estimate 3.04 * 5.3" and given choices of 1.6, 16, 160, and 1600, “[t]he 13-year olds' responses were almost at the level of guessing, and the performance by 17-year-olds was not much better.” Students do not understand what multiplication means, even though they can perform the procedure.

This problem has been studied for almost a century. In the 1920s, Louis Benezet, who had just become superintendent of schools in Manchester, New Hampshire, noted that the traditional teaching of arithmetic was “chloroform[ing] the child's reasoning faculties.”[10] As an example, he gave students in a traditionally taught classroom the following problem:

7 V. V. Davydov (1990), Types of Generalization in Instruction: Logical and Psychological Problems in the Structuring of School Curricula. Soviet Studies in Mathematics Education. Volume 2. National Council of Teachers of Mathematics (original work published in 1972 in Russian). See also Yuriy V. Karpov (1995), "L. S. Vygotsky as the founder of a new approach to instruction.," School Psychology International 16(2):131-142.

8 P. Ia. Galperin and N. S. Pantina (1965), "The Dependency Of The Motor Habit On The Type Of The Task Orientation," Russian Monographs on Brain and Behavior 3:425-433 (original work published in 1957 in Russian).

9 Thomas P. Carpenter, et al. (1983), "Results of the third NAEP mathematics assessment: Secondary school," The Mathematics Teacher 76(9):652-659.

10 Louis P. Benezet (1935, 1936), “The teaching of arithmetic I, II, III: The story of an experiment," Journal of the National Education Association 24(8):241-244 (I, Nov 1935), 24(9):301-303 (II, Dec 1935), 25(1):7-8 (III, Jan 1936), available at the Benezet Centre. The "chloroform" quote is from part I. For a study of Benezet's curriculum, showing very positive results, see Etta Berman (1935), The Result of Deferring Systematic Teaching of Arithmetic to Grade Six as Disclosed by the Deferred Formal Arithmetic Plan at Manchester, New Hampshire, Masters Thesis, Boston University, USA.

Three steamers leave Boston simultaneously for Portland [Maine]. One takes 12 hours; the other takes 15 hours; and the third takes 17 hours. How long before all the steamers are in Portland?

In the ninth grade (roughly age 14), only 6 out of 29 students solved the problem correctly. (As a result of such performances, Benezet instituted a mathematics curriculum in grades 1-7 that emphasized understanding and using mathematics, with the procedural aspects mastered in a few months in the 7th grade.)

The problem of lacking meaning is worldwide and is separate from computational fluency. An example comes from the famous NAEP school-bus problem:[11]

1128 soldiers have to be bused to a base; each bus holds 36 soldiers. How many bus are needed?

Even though the correct answer is 32, and even though most students can set up the division correctly, the most common answer by US students (on the NAEP) was the meaningless "31 R 12" (31 with a remainder of 12). Unfortunately, the failure to make meaning of mathematics affects students in China as well, despite the greater computational fluency of Chinese students.[12]

Verschaffel and colleagues, studying students in Belgium, Japan, and beyond, found that students are reluctant to use real-world information to solve or even make sense of mathematics problems.[13] For example, when asked for the water temperature after combining 40- and 120-degree water, students answer with 160 degrees---an answer divorced from any meaning. In contrast, if you ask students to describe the result of mixing hot and cold water---that is, if the question doesn't directly involve mathematics---they correctly say warm or lukewarm. When the problem becomes mathematical, thinking too often gives way to rote procedures.

The good news is that testing drives education

Rote learning is so fundamental and widespread that one might despair. Fortunately, assessment drives learning and teaching. That observation, described extensively in medical education,[14] provides the lever to guide mathematics teaching toward 21st-century skills. The "students" of PISA are the OECD and associated countries. When PISA sets a high standard, countries follow in their teaching practices. The four PISA contexts, developed in the 2012 assessment, have solved the problem of “What should mathematics be for?” If, in each area, the measure is understanding, countries will shift their education away from rote learning and toward real understanding.

Test for understanding

11 Carpenter (1983), supra.

12 Jinfa Cai and Edward A. Silver (1995), "Solution processes and interpretations of solutions in solving a division-with-remainder story problem: Do Chinese and US students have similar difficulties?" Journal for Research in Mathematics Education 26(5):491-497.

13 Lieven Verschaffel, Brian Greer, and Erik de Corte (2000), Making Sense of Word Problems, Swets and Zeitlinger; Hajime Yoshida, Lieven Verschaffel, and Erik de Corte (1997), "Realistic considerations in solving problematic word problems: Do Japanese and Belgian children have the same difficulties?" Learning and Instruction 7(4):329--338.

14 D. I. Newble and Kerry Jaeger (1983), "The effect of assessments and examinations on the learning of medical students," Medical Education 17(3):165-71. See also Marjolein Heijne-Penninga (2008), "Influence of open- and closed-book tests on medical students' learning approaches," Medical Education 42(10):967-974.

Thus, the marking of questions needs to have the correct sign, in that answers, or justifications, that reveal lack of understanding “contaminate the whole answer.”[15] Vinner has called this kind of rote thinking “pseudo-conceptual” or “pseudo-analytic” because it contains only a facsimile of analysis or conceptual knowledge.[16] Being able to detect whether analysis is real or fake requires the right kind of question, one where rote learning and understanding lead to different conclusions, or where rote learning fails entirely. As Vinner illustrates, if the question is, “A rectangle has side lengths 7 and 5 cm. What is its area?” then understanding or rote learning will lead to the answer of 35 cm^2. However, if the question is, “The perimeter of a rectangle is 24 cm, and one side has length 7 cm. What is the rectangle's area?” then the rote approach will often give 7 * 24 = 168, and the student will write down 168 cm^2.

An example from PISA is the "Climbing Mount Fuji" unit.[17] The question asked students to find the step length of a person who walked 22,500 steps up Mount Fuji covering 9 km. A small proportion of students got partial credit for giving 0.4 cm or 4000 cm as answers. The 0.4 cm can be understood as meaning meters and forgetting to include the units. But 4000 cm has to be nonsense, whether as meters of centimeters. That it could be written down shows that the student has no conception of mathematics as describing the world.

Similarly, the "Pizzas" unit[18] allows students to calculate the answer without understanding proportional reasoning. They can mindlessly compute the area of the pizzas, divide the areas by the respective cost, and compare the results---a very low-level skill. In contrast, if the pizzas share the same irregular shape such as a kidney bean (at different sizes), then proportional reasoning is the only way to solve the problem.

With a repertoire of questions that distinguish rote learning from understanding, the marking scales in PISA could reliably emphasize measuring what students understand about mathematics as a tool for making sense of the world.

Mathematics for transfer in the 21st century

The mathematics that students need for full participation in society, in all four contexts, can be divided into reasoning tools. The tools fall into two groups: tools that build connection among quantities (and, more generally, mathematical objects) and tools that simplify the world.

(A) Tools for connection
- (1) comparison
- (2) proportional reasoning
- (3) multiplicative scales (including exponential growth)
- (4) probabilistic (including logical) reasoning

(B) Tools for simplification

---

15 Eric M. Rogers (1969), "Examinations: Powerful agents for good or ill in teaching," American Journal of Physics 37:954-962.

16 Shlomo Vinner (1997), "The pseudo-conceptual and the pseudo-analytical thought processes in mathematics learning," Educational Studies in Mathematics 34:97–129. Similarly, Skemp distinguishes between instrumental and relational understanding [Richard Skemp (1976), "Relational Understanding and Instrumental Understanding," Mathematics Teaching 77:20-26]. In contrast to instrumental understanding (the analog of rote learning), relational understanding promotes itself, and leads to more relational understanding, as the network of understanding grows and thickens.

17 PISA 2012 Assessment and Analytical Framework, p. 48 (p. 49 in the PDF file).

18 PISA 2012 Assessment and Analytical Framework, pp. 49-51 (pp. 50-52 in the PDF file).

(5) divide and conquer
(6) lumping
(7) easy cases

Tools for connection

Man, said the Greeks, is a social animal: Without any conection to others, we are not fully human, and life loses meaning. Quantities are similar. Alone, a quantity has no meaning; it is just an isolated fact. Meaning comes from connections to other quantities. In contrast to computers, we humans are meaning-making creatures. We need reasoning tools to build connection. Thus, this first set of tools---tools for connection---is a first major step toward using mathematics to represent the complexity of the world.

(1) *comparison*. No quantity is meaningful by itself, and thus must be compared to another, related quantity to give it meaning. For example, a monetary rate of $1 trillion dollars per year means little by itself. As a fraction of a government budget or a GDP, it acquires meaning for us. The essential skill is making meaningful dimensionless ratios.

Comparison a fundamental part of critical thinking.[19] For example, reading a bare number in a newspaper article, such as "1 million fans came into the streets to celebrate their sports team" (as appears regularly in Boston newspapers) conveys little information. Only when the 1 million is compared to the total population of Boston and the inner suburbs, also approximately 1 million, does the number acquire meaning (as being highly implausible).

(2) *proportional reasoning*. Comparison naturally produces ratios. Understanding how ratio changes are related (e.g. doubling lengths quadruples areas)---the essence of proportional reasoning---is essential for engineering and science. For example, the quadratic proportionality between diffusion times and distance means that large animals must have a circulatory system.

Proportional reasoning, explained the renowned mathematics educator Hans Freudenthal, deserves “priority above algorithmic computations and applications of formulae because it deepens the insight and the rich context in the naive, scientific, and social reality where it operates.”[20] Without proportional reasoning, the world makes no sense. With it, we find the world rich in quantities connected through ratio. Seeing the world through ratio is hardwired into our brains' approximate number system.[21] Our teaching should build on this inherent and powerful ability to make sense of the world.

(3) *multiplicative scales (including exponential growth)*. Once ratio is seen as essential, the next step is counting ratios---e.g. by how many factors of 2 (or 10) is the sun larger in mass than the earth? Counting ratios leads to logarithmic and exponential scales. They are equally essential for the social and physical worlds. Logarithmic scales organize the huge dynamic range of lengths, times, energies in the universe. In the social world, exponential-growth models help understand the growth of populations, economies, or resource use. Without understanding

[19] Critical thinking is one of the 4 C's in Bernie Trilling and Charles Fadel (2009), 21st Century Skills: Learning for Life in Our Times, Jossey-Bass.

[20] Hans Freudenthal (1983), Didactical Phenomenology of Mathematical Structures, Reidel, p. 401.

[21] Stanislas Dehaene (2011), The Number Sense: How the Mind Creates Mathematics, Oxford University Press.

and taming exponential growth, the world has no chance of using its resources wisely.[22] Proportional reasoning and exponential growth need to be understood at a gut level, simply as part of how students perceive the world quantitatively. Both ideas take a short time to describe, but a long time to master.[23]

(This mastery would come much quicker if our teaching of ratio, proportionality, and logarithmic scales were aligned with our innate perception of quantity. Instead, students are untaught our innate multiplicative scale, and then retaught it much later as a series of seemingly arbitrary and therefore meaningless rules for manipulating logarithms. Teachers of physical science and engineering can list numerous examples of students' difficulties with logarithms.)

Within the scope of this tool falls the Rule of 72 for estimating the doubling time of an exponential process given the growth rate.[24] As an illustrative assessment, though one unlikely to be adopted by many governments, students (without a calculator) could be asked, "Your national government plans to underestimate, by 2 percent per year, the rate of inflation. If you will retire in 40 years, by roughly what factor will your public pension be devalued?"

(4) *probabilistic (including logical) reasoning*. Our knowledge of the world is necessarily incomplete. We represent our incomplete knowledge through probability---in particular, the probabilities of hypotheses. Collecting data and evidence changes our knowledge and thus our probabilities according to Bayes theorem.[25] Bayes theorem connects evidence and belief. It is a tool for connection.

If students understand Bayes theorem, they do not fall into the prosecutor's fallacy or the defense-attorney's fallacy.[26] They can judge the strength of clinical evidence, and acquire a way of thinking that matches strength of conviction to evidence. This tool is essential for critical thinking.

Probabilistic reasoning contains logical reasoning: In the extreme case where probabilities are either 0 (false) or 1 (true), probabilistic reasoning simplifies to propositional logic (logical reasoning). The best example of logical reasoning taught for transfer is Harold Fawcett's geometry course from the 1930s.[27] Instead of of reproducing geometric constructions, which are anyway performed more precisely by computers, students learned proof in order to

22 Albert A. Bartlett (2004), The Essential Exponential! For the Future of Our Planet, Center for Science, Mathematics and Computer Education, University of Nebraska, Lincoln. Al Bartlett's lecture on the importance of understanding exponential growth has over 5 million views on YouTube. The subsequent seven parts each have more than 500,000 views.

23 On the difficulties student face in mastering proportional reasoning, see for example Dirk de Bock et al. (2002), "Improper use of linear reasoning: An in-depth study of the nature and the irresistibility of secondary school students' errors," Educational Studies in Mathematics 50(3): 311-334. On the long time to master multiplicative structures, see OECD (2009), Learning Math for Life: A Perspective from PISA, p. 177 (p.179 of the PDF file).

24 Alberta Education (2008), The Alberta 10--12 mathematics program of studies with achievement indicators, Alberta, Canada, p. 92.

25 Edwin Jaynes (2003), Probability Theory: The Logic of Science, Cambridge University Press.

26 William C. Thompson and Edward L. Schumann (1987), "Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy," Law and Human Behavior 2(3):167-187. See also Leila Schneps and Coralie Colmez (2013), Math on Trial: How Numbers Get Used and Abused in the Courtroom, Basic Books.

27 Harold Fawcett (1938), The Nature of Proof, Bureau of Publications, Columbia University. Reprinted by the NCTM in 1995 and 2001. See also Erica Lane (2004), "The Nature of Proof in Today's Classroom," The Montana Mathematics Enthusiast (TMME) 1(2):58-65; and Fred Flener (2001, April 6), "A geometry course that changed their lives: The guinea pigs after 60 years," Annual Conference of the NCTM, Orlando, Florida.

deconstruct advertising copy or improve the framing of laws---true critical thinking. An illustrative problem from the course was the following:

While reading a magazine Helen's attention was drawn to the picture of a beautiful girl with an attractive smile who was represented as saying: “I'd wished a thousand times for a brighter smile. One tube of Colgate's gave it to me. It was so annoying to see other girls with lovely smiles get all the dates. Then I tried Colgate's. Now my smiles are bright too.”

This is really an argument for using Colgate's Dental Cream. The argument is based on certain assumptions. What are these assumptions?

It is hard to imagine a more powerful or transferable ability.

Tools for simplification

The world is filled with quantities and other mathematical objects. Thus, giving them meaning, which is the purpose of the first group of tools, can be dangerous. The meanings will overwhelm us. Thus, we cannot shut our toolbox yet. The second group of tools---tools for simplification---enable us to fit the complexity of the world into our limited minds.

(5) *divide and conquer*. We often split hard problems into manageable pieces. However, this essential problem-solving skill needs naming, so that students realize when they are using it, and have a method to begin even difficult problems. This tool is valuable from the earliest ages. For example, in adding 24 + 38, students can divide 38 into 6 and 32, and then add 24+6=30 and 30+32=62.

For PISA-age students, a class of problems that use divide-and-conquer reasoning is: "At what rate (passengers per hour) can a busy train track carry passengers?" Students must divide the problem into simpler estimates: how many trains per hour, how many cars on each train, and how many passengers in each train. (This problem also tests their ability to be comfortable with imprecision, an ability developed by the next reasoning tool of lumping.)

This tool is essential to solving interview questions for management-consulting positions at leading firms such as McKinsey. The task of a consultant is to understand quickly the essential attributions of a problem. Skills essential for consulting are essential in the occupational context.

(6) *lumping*. Rigor leads to rigor mortis. When we round numbers or approximate complicated shapes with simple shapes, we throw away less important details, in order to free our minds to master the essential parts of a problem. Only in this way can be master complexity. A student who has mastered lumping could never answer that 3.04 * 5.3 is approximately 1600.

(7) *easy cases*. When a problem is still too hard, we look at its special cases, ones that we can understand. For example, in evaluating a mathematical conjecture, we first test it with n=0 and n=1. Or, in the personal and societal contexts, mortgages can, at one extreme, be like annuities:

interest rate * loan period >> 1

At the other extreme, they can like installment loans:

interest rate * loan period << 1

In each extreme the payment is much easier to calculate, and the calculation is much easier to understand, than in the general case. Thinking about easy cases promotes insight.

This kind of reasoning is particularly valuable in the scientific context. For example, in Liping Ma's study of American and Chinese mathematics teachers,[28] teachers were asked to propose answers to a student who says, "I discovered that the bigger the perimeter of a rectangle, the bigger the area." Many teachers did not know how to decide whether the claim was true. The method of easy cases provides an insightful route: Make the rectangle thinner but longer. In that way, the perimeter can be made as large and the area can be made as close to zero as one wants. Thus, the conjecture must be false. If teachers and, more importantly, their students can apply this reasoning, they gain a valuable tool for testing and developing new conjectures.

Conclusion: Reasoning tools fight rote learning and promote transfer

Reasoning tools, by their nature, transfer across domains. They are therefore a powerful antidote to rote learning. By emphasizing real understanding, PISA's assessment will push the mathematics curriculum worldwide toward competence in the four PISA contexts: personal, societal, occupational, and scientific. By offering reasoning tools as the organizing curriculum principle, PISA can help countries reach this promised land.

---

[28] Liping Ma (1999), Knowing and Teaching Elementary Mathematics: Teachers' Understanding of Fundamental Mathematics in China and the United States, Lawrence Erlbaum Associates.